

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 15/57, 9/64, 9/54, 1/21, A61K 38/48</b>		<b>A2</b>	(11) International Publication Number: <b>WO 99/53078</b>
			(43) International Publication Date: 21 October 1999 (21.10.99)
(21) International Application Number: <b>PCT/US99/08177</b>			(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
(22) International Filing Date: 15 April 1999 (15.04.99)			
(30) Priority Data: 09/060,854 15 April 1998 (15.04.98) US			
(71) Applicant: GENENCOR INTERNATIONAL, INC. [US/US]; 4 Cambridge Place, 1870 South Winton Road, Rochester, NY 14618 (US).			
(72) Inventor: ESTELL, David; 248 Woodbridge Circle, San Mateo, CA 94403 (US).			
(74) Agent: STONE, Christopher, L.; Genencor International, Inc., 925 Page Mill Road, Palo Alto, CA 94304-1013 (US).			<b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: HUMAN PROTEASE AND USE OF SUCH PROTEASE FOR PHARMACEUTICAL APPLICATIONS AND FOR REDUCING THE ALLERGENICITY OF NON-HUMAN PROTEINS			
(57) Abstract  The present invention relates to a method of producing novel improved protein mutant which produce low allergenic response in humans compared to the parent of that mutant. Specifically, the present invention comprises neutralizing or reducing the allergenicity of a protein by introducing therein as replacement or modification of an epitope on such protein a sequence from human subtilisin.			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

**HUMAN PROTEASE AND USE OF SUCH PROTEASE FOR  
PHARMACEUTICAL APPLICATIONS AND FOR REDUCING THE ALLERGENICITY OF  
NON-HUMAN PROTEINS**

5

**BACKGROUND OF THE INVENTION**

**A. Field of the Invention**

The present invention relates to a human protein sequence which can be used in several applications. Specifically, the novel human protein sequence can be used to design proteins which produce lower allergenic response in humans exposed to such proteins through the use of a predictive assay.

**B. State of the Art**

Serine proteases are a subgroup of carbonyl hydrolases. They comprise a diverse class of enzymes having a wide range of specificities and biological functions. Stroud, R. Sci. Amer., 131:74-88. Despite their functional diversity, the catalytic machinery of serine proteases has been approached by at least two genetically distinct families of enzymes: the subtilisins and the mammalian chymotrypsin related and homologous bacterial serine proteases (e.g., trypsin and *S. gresius* trypsin). These two families of serine proteases show remarkably similar mechanisms of catalysis. Kraut, J. (1977), Ann. Rev. Biochem., 46:331-358. Furthermore, although the primary structure is unrelated, the tertiary structure of these two enzyme families bring together a conserved catalytic triad of amino acids consisting of serine, histidine and aspartate.

Subtilisin is a serine endoprotease (MW 27,500) which is secreted in large amounts from a wide variety of *Bacillus* species and other microorganisms. The protein sequence of subtilisin has been determined from at least four different species of *Bacillus*. Markland, F.S., et al. (1983), Honne-Seyler's Z. Physiol. Chem., 364:1537-1540. The three-dimensional crystallographic structure of *Bacillus amyloliquefaciens* subtilisin to 2.5Å resolution has also been reported. Wright, C.S., et al. (1969), Nature, 221:235-242; Drenth, J., et al. (1972), Eur. J. Biochem., 26:177-181. These studies indicate that although subtilisin is genetically unrelated to the mammalian chymotrypsin like serine proteases, it has a similar active site structure. The x-ray crystal structures of subtilisin containing covalently bound peptide inhibitors (Robertus, J.D., et al. (1972), Biochemistry, 11:2439-2449) or product complexes (Robertus, J.D., et al. (1976), J. Biol. Chem., 251:1097-1103) have also provided information regarding the active site and putative substrate binding cleft of subtilisin. In addition, a large number of kinetic and chemical

modification studies have been reported for subtilisin (Philipp, M., et al. (1983), Mol. Cell. Biochem., 51:5-32; Svendsen, B. (1976), Carlsberg Res. Comm., 41:237-291; Markland, F.S. Id.) as well as at least one report wherein the side chain of methionine at residue 222 of subtilisin was converted by hydrogen peroxide to methionine-sulfoxide (Stauffer, D.C., et al. (1965), J. Biol. Chem., 244:5333-5338) and the side chain of serine at residue 221 converted to cysteine by chemical modification (Polgar, et al. (1981), Biochimica et Biophysica Acta, 667:351-354.)

Proteins bearing some resemblance and/or homology to bacterial subtilisin have also been detected in humans as well (see e.g., Keifer et al., DNA and Cell Biol., Vol. 10, No. 10, pp. 757-769 (1991); Smeekens et al., J. Biol. Chem., Vol. 265, No. 6, pp. 2997-3000 (1990); Tomkinson et al., Biochem., Vol. 30, pp. 168-174 (1991)).

US Patent 4,760,025 (RE 34,606) discloses the modification of subtilisin amino acid residues corresponding to positions in *Bacillus amyloliquefaciens* subtilisin tyrosine - 1, aspartate +32, asparagine +155, tyrosine +104, methionine +222, glycine +166, histidine +64, glycine +169, phenylalanine +189, serine +33, serine +221, tyrosine +217, glutamate +156 and alanine +152. US Patent 5,182,204 discloses the modification of the amino acid +224 residue in *Bacillus amyloliquefaciens* subtilisin and equivalent positions in other subtilisins which may be modified by way of substitution, insertion or deletion and which may be combined with modifications to the residues identified in US Patent 4,760,025 (RE 34,606) to form useful subtilisin mutants or variants. US Patent 5,155,033 discloses similar mutant subtilisins having a modification at an equivalent position to +225 of *B. amyloliquefaciens* subtilisin. US Patents 5,185,258 and 5,204,015 disclose mutant subtilisins having a modification at positions +123 and/or +274. US Patent 5,182,204 discloses the modification of many amino acid residues within subtilisin, including specifically +99, +101, +103, +107, +126, +128, +135, +197 and +204. US Patent 4,914,031 discloses certain subtilisin analogs, including a subtilisin modified at position +76.

Proteins, including proteases, used in industrial, pharmaceutical and commercial applications are of increasing prevalence. As a result, the increased exposure due to this prevalence has been responsible for some safety hazards caused by the sensitization of certain persons to those peptides, whereupon subsequent exposure causes extreme allergic reactions which can be injurious and even fatal. For example, proteases are known to cause dangerous hypersensitivity in some individuals. As a result, despite the usefulness of proteases in industry, e.g., in laundry detergents, cosmetics, textile treatment etc., and the extensive research performed in the field to provide improved

proteases which have, for example, more effective stain removal under detergency conditions, the use of proteases in industry has been problematic due to their ability to produce a hypersensitive allergic response in some humans.

Much work has been done to alleviate these problems. Among the strategies explored to reduce immunogenic potential of protease use have been improved production processes which reduce potential contact by controlling and minimizing workplace concentrations of dust particles or aerosol carrying airborne protease, improved granulation processes which reduce the amount of dust or aerosol actually produced from the protease product, and improved recovery processes to reduce the level of potentially allergenic contaminants in the final product. However, efforts to reduce the allergenicity of protease, per se, have been relatively unsuccessful. Alternatively, efforts have been made to mask epitopes in protease which are recognized by immunoglobulin E (IgE) in hypersensitive individuals (PCT Publication No. WO 92/10755) or to enlarge or change the nature of the antigenic determinants by attaching polymers or peptides/proteins to the problematic protease.

When an adaptive immune response occurs in an exaggerated or inappropriate form, the individual experiencing the reaction is said to be hypersensitive. Hypersensitivity reactions are the result of normally beneficial immune responses acting inappropriately and sometimes cause inflammatory reactions and tissue damage. They can be provoked by many antigens; and the cause of a hypersensitivity reaction will vary from one individual to the next. Hypersensitivity does not normally manifest itself upon first contact with the antigen, but usually appears upon subsequent contact. One form of hypersensitivity occurs when an IgE response is directed against innocuous environmental antigens, such as pollen, dust-mites or animal dander. The resulting release of pharmacological mediators by IgE-sensitized mast cells produces an acute inflammatory reaction with symptoms such as asthma or rhinitis.

Nonetheless, a strategy comprising modifying the IgE sites will not generally be successful in preventing the cause of the initial sensitization reaction. Accordingly, such strategies, while perhaps neutralizing or reducing the severity of the subsequent hypersensitivity reaction, will not reduce the number or persons actually sensitized. For example, when a person is known to be hypersensitive to a certain antigen, the general, and only safe, manner of dealing with such a situation is to isolate the hypersensitive person from the antigen as completely as possible. Indeed, any other course of action would be dangerous to the health of the hypersensitive individual. Thus, while reducing

the danger of a specific protein for a hypersensitive individual is important, for industrial purposes it would be far more valuable to render a protein incapable of initiating the hypersensitivity reaction in the first place.

- 5 T-lymphocytes (T-cells) are key players in the induction and regulation of immune responses and in the execution of immunological effector functions. Specific immunity against infectious agents and tumors is known to be dependent on these cells and they are believed to contribute to the healing of injuries. On the other hand, failure to control these responses can lead to auto aggression. In general, antigen is presented to T-cells
- 10 in the form of antigen presenting cells which, through a variety of cell surface mechanisms, capture and display antigen or partial antigen in a manner suitable for antigen recognition by the T-cell. Upon recognition of a specific epitope by the receptors on the surface of the T-cells (T-cell receptors), the T-cells begin a series of complex interactions, including proliferation, which result in the production of antibody by B-cells.
- 15 While T-cells and B-cells are both activated by antigenic epitopes which exist on a given protein or peptide, the actual epitopes recognized by these mononuclear cells are generally not identical. In fact, the epitope which activates a T-cell to initiate the creation of immunologic diversity is quite often not the same epitope which is later recognized by B-cells in the course of the immunologic response. Thus, with respect to hypersensitivity,
- 20 while the specific antigenic interaction between the T-cell and the antigen is a critical element in the initiation of the immune response to antigenic exposure, the specifics of that interaction, i.e., the epitope recognized, is often not relevant to subsequent development of a full blown allergic reaction.

PCT Publication No. WO 96/40791 discloses a process for producing polyalkylene oxide-polypeptide conjugates with reduced allergenicity using polyalkylene oxide as a

25 starting material.

PCT Publication No. WO 97/30148 discloses a polypeptide conjugate with reduced allergenicity which comprises one polymeric carrier molecule having two or more polypeptide molecules coupled covalently thereto.

30 PCT Publication No. WO 96/17929 discloses a process for producing polypeptides with reduced allergenicity comprising the step of conjugating from 1 to 30 polymolecules to a parent polypeptide.

PCT Publication No. WO 92/10755 discloses a method of producing protein variants evoking a reduced immunogenic response in animals. In this application, the

35 proteins of interest, a series of proteases and variants thereof, were used to immunized

rats. The sera from the rats was then used to measure the reactivity of the polyclonal antibodies already produced and present in the immunized sera to the protein of interest and variants thereof. From these results, it was possible to determine whether the antibodies in the preparation were comparatively more or less reactive with the protein and its variants, thus permitting an analysis of which changes in the protein are likely to neutralize or reduce the ability of the Ig to bind. From these tests on rats, the conclusion was arrived at that changing any of subtilisin 309 residues corresponding to 127, 128, 129, 130, 131, 151, 136, 151, 152, 153, 154, 161, 162, 163, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 186, 193, 194, 195, 196, 197, 247, 251, 261 will result in a change in the immunological potential.

PCT Publication No. WO 94/10191 discloses low allergenic proteins comprising oligomeric forms of the parent monomeric protein, wherein the oligomer has substantially retained its activity.

The prior art has provided methods of reducing the allergenicity of certain proteins and identification of epitopes which cause allergic reactions in some individuals, the assays used to identify these epitopes generally involving measurement of IgE and IgG antibody in blood sera previously exposed to the antigen. Nonetheless, a need continues for alternate methods of preparing low allergenicity enzymes. Likewise, a need exists for an increased availability of human enzymes which may have use in pharmaceutical applications.

### **SUMMARY OF THE INVENTION**

It is an object of the invention to provide a human protease which can be used in industry as a replacement for bacterial and fungal proteases.

It is an object of the invention to provide a method of making currently used and successful proteases and other proteins more safe by integrating therein sequences derived from human protease analogs.

It is a further object of the invention to provide a human protease which may have application in the pharmaceutical industry.

According to the present invention, a method for reducing the allergenicity of a non-human protein is provided wherein an epitope is identified and replaced with an analogous region within a human subtilisin. In a preferred embodiment the non-human protein is an enzyme, more preferably a protease. In another preferred embodiment, the epitope replaced is a T-cell epitope.

In another embodiment of the present invention, a method for producing the protein of the invention having reduced allergenicity is provided. Preferably, the mutant protein is prepared by modifying a DNA encoding a precursor protein so that the modified DNA encodes the mutant protein of the invention wherein an epitope is replaced with an analogous region from human subtilisin.

In yet another embodiment of the invention, DNA sequences encoding the mutant protein, as well as expression vectors containing such DNA sequences and host cells transformed with such vectors are provided, which host cells are preferably capable of expressing such DNA to produce the mutant protein of the invention either intracellularly or extracellularly.

The mutant protein of the invention is useful in any composition or process in which the protein is generally known to be useful. For example, where the protein is a protease, the reduced allergenicity protease can be used as a component in cleaning products such as laundry detergents and hard surface cleansers, as an aid in the preparation of leather, in the treatment of textiles such as wool and/or silk to reduce felting, as a component in a cosmetic or face cream, and as a component in animal or pet feed to improve the nutritional value of the feed. Similarly, where the protein is an amylase, the reduced allergenicity amylase can be used for the liquefaction of starch, as a component in a dishwashing detergent, for desizing of textiles, in a laundry detergent or any other use for which amylase is useful. Similarly, where the protein is a pharmaceutical composition, its use can be made more safe by reducing the possibility of allergic reaction.

In another embodiment of the invention, the human subtilisin may be used in pharmaceutical applications wherein the protease is used for debridement treatments.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

Figs. 1 A-C illustrates the DNA (SEQ ID:NO 1) and amino acid (SEQ ID:NO 2) sequence for *Bacillus amyloliquefaciens* subtilisin (BPN') and a partial restriction map of this gene.

Fig. 2 illustrates the conserved amino acid residues among subtilisins from *Bacillus amyloliquefaciens* and *Bacillus lentus* (wild-type)(SEQ ID:NO 3).

Figs. 3A and 3B illustrate an amino acid sequence alignment of subtilisin type proteases from *Bacillus amyloliquefaciens* (BPN'), *Bacillus subtilis*, *Bacillus licheniformis* (SEQ ID:NO 4) and *Bacillus lentus*. The symbol \* denotes the absence of specific amino acid residues as compared to subtilisin BPN'.



Fig. 4. illustrates the additive T-cell response of 16 peripheral mononuclear blood samples to peptides corresponding to the *Bacillus lentus* protease. Peptide E05 represents the region comprising residues corresponding to 170-173 in protease from *Bacillus amyloliquefaciens*.

5            Fig. 5 illustrate the additive T-cell response of 10 peripheral mononuclear blood sample to peptides corresponding to the human subtilisin.

Fig. 6 illustrates the amino acid sequence of human subtilisin (SEQ ID:NO 6).

Fig. 7 illustrates the amino acid strings corresponding to peptides derived from the sequence of *Bacillus lentus* protease used in Example 2.

10           Figs. 8A and 8B illustrate the amino acid strings corresponding to peptides derived from the sequence of human subtilisin used in Example 2.

#### **DETAILED DESCRIPTION OF THE INVENTION**

According to the present invention, a method for reducing the allergenicity of a non-human protein is provided wherein an epitope is identified and replaced with an analogous region within a human subtilisin. In a preferred embodiment the non-human protein is an enzyme, more preferably a protease. In another preferred embodiment, the epitope replaced is a T-cell epitope.

15           In another embodiment of the present invention, a method for producing the protein of the invention having reduced allergenicity is provided. Preferably, the mutant protein is prepared by modifying a DNA encoding a precursor protein so that the modified DNA encodes the mutant protein of the invention wherein an epitope is replaced with an analogous region from human subtilisin.

20           In yet another embodiment of the invention, DNA sequences encoding the mutant protein, as well as expression vectors containing such DNA sequences and host cells transformed with such vectors are provided, which host cells are preferably capable of expressing such DNA to produce the mutant protein of the invention either intracellularly or extracellularly.

25           According to a preferred embodiment of the present invention, the epitope to be replaced in the non-human protein of interest is identified by a method for identifying T-cell epitopes. In a preferred embodiment of the invention, the present invention provides an assay which identifies epitopes as follows: differentiated dendritic cells are combined with naïve human CD4+ and/or CD8+ T-cells and with a peptide of interest. More specifically, a method is provided wherein a T-cell epitope is recognized comprising the steps of: (a) obtaining from a single blood source a solution of dendritic cells and a solution of naïve

30  
35

CD4+ and/or CD8+ T-cells; (b) promoting differentiation in said solution of dendritic cells; (c) combining said solution of differentiated dendritic cells and said naïve CD4+ and/or CD8+ T-cells with a peptide of interest; (d) measuring the proliferation of T-cells in said step (c).

5           The non-human peptide of interest to be analyzed according to the assay of the invention is derived from a protein or enzyme for which reduced allergenicity is required. In the practice of the invention, it is possible to identify with precision the location of an epitope which can cause sensitization in an individual or sampling of individuals. In a particularly effective embodiment of the invention, a series of peptide oligomers which  
10       correspond to all or part of the protein or enzyme are prepared. For example, a peptide library is produced covering the relevant portion or all of the protein. One particularly useful manner of producing the peptides is to introduce overlap into the peptide library, for example, producing a first peptide corresponds to amino acid sequence 1-10 of the subject protein, a second peptide corresponds to amino acid sequence 4-14 of the subject  
15       protein, a third peptide corresponds to amino acid sequence 7-17 of the subject protein, a fourth peptide corresponds to amino acid sequence 10-20 of the subject protein etc. . . . until representative peptides corresponding to the entire molecule are created. By analyzing each of the peptides individually in the assay provided herein, it is possible to precisely identify the location of epitopes recognized by T-cells. In the example above, the  
20       reaction of one specific peptide to a greater extent than its neighbors will facilitate identification of the epitope anchor region to within three amino acids. After determining the location of these epitopes, it is possible to alter the amino acids within each epitope until the peptide produces a less significant T-cell response.

          Preferably, the epitope is modified in one of the following ways: (a) preferably the  
25       amino acid sequence of the epitope is substituted with an analogous sequence from the human subtilisin of the invention to the protein of interest, e.g., where the protein is a subtilisin, a sequence alignment can be arranged so as to find the analogous region in the human subtilisin molecule with which to replace the pertinent epitope in the subtilisin; (b) the amino acid sequence of the epitope is substituted with a sequence from human  
30       subtilisin of the invention which substantially mimics the major tertiary structure attributes of the epitope, but which produces a lesser allergenic response due to T-cell epitope recognition than that of the protein of interest; or (c) with any sequence from the human subtilisin of the invention which produces lesser allergenic response due to T-cell epitope recognition than that of the protein of interest.

"Antigen presenting cell" as used herein means a cell of the immune system which present antigen on their surface which is recognizable by T-cells. Examples of antigen presenting cells are dendritic cells, interdigitating cells, activated B-cells and macrophages.

5 "T-cell proliferation" as used herein means the number of T-cells produced during the incubation of T-cells with the antigen presenting cells, with or without antigen.

"Baseline T-cell proliferation" as used herein means T-cell proliferation which is normally seen in an individual in response to exposure to antigen presenting cells in the absence of peptide or protein antigen. For the purposes herein, the baseline T-cell  
10 proliferation level was determined on a per sample basis for each individual as the proliferation of T-cells in response to antigen presenting cells in the absence of antigen.

"T-cell epitope" means a feature of a peptide or protein which is recognized by a T-cell receptor in the initiation of an immunologic response to the peptide comprising that antigen. Recognition of a T-cell epitope by a T-cell is generally believed to be via a  
15 mechanism wherein T-cells recognize peptide fragments of antigens which are bound to class I or class II major histocompatibility (MHC) molecules expressed on antigen-presenting cells (see e.g., Moeller, G. ed., Antigenic Requirements for Activation of MHC-Restricted Responses, Immunological Review, Volume 98, p 187 (Copenhagen; Munksgaard) (1987).

20 The epitopes determined according to the assay provided herein are then modified to reduce the allergenic potential of the protein of interest. In a preferred embodiment, the epitope to be modified produces a level of T-cell proliferation of greater than three times the baseline T-cell proliferation in a sample. When modified, the epitope produces less than three times the baseline proliferation, preferably less than two times the baseline  
25 proliferation and most preferably less than or substantially equal to the baseline proliferation in a sample.

"Sample" as used herein comprises mononuclear cells which are naïve, i.e., not sensitized, to the antigen in question.

"Homolog" as used herein means a protein or enzyme which has similar catalytic  
30 action, structure and/or use as the protein of interest. It is desirable to find a homolog that has a tertiary and/or primary structure similar to the protein of interest as replacement of the epitope in the protein of interest with an analogous segment from the homolog will reduce the disruptiveness of the change. Thus, enzymes having significant homology will provide the most desirable target for epitope substitutions with sequences from the human  
35 subtilisin of the invention.

An "analogous" sequence may be determined by ensuring that the replacement amino acids show a similar function, the tertiary structure and/or conserved residues to the amino acids in the protein of interest at or near the epitope. Thus, where the epitope region contains, for example, an alpha-helix or a beta-sheet structure, the replacement amino acids should maintain that specific structure.

While the present invention extends to all proteins for which it is desired to reduce allergenicity, for the sake of simplicity, the following will describe a particularly preferred embodiment of the invention, the modification of protease. Proteases are carbonyl hydrolases which generally act to cleave peptide bonds of proteins or peptides. As used herein, "protease" means a naturally-occurring protease or a recombinant protease. Naturally-occurring proteases include  $\alpha$ -aminoacylpeptide hydrolase, peptidylamino acid hydrolase, acylamino hydrolase, serine carboxypeptidase, metallocarboxypeptidase, thiol proteinase, carboxylproteinase and metalloproteinase. Serine, metallo, thiol and acid proteases are included, as well as endo and exo-proteases.

Subtilisins are bacterial or fungal proteases which generally act to cleave peptide bonds of proteins or peptides. As used herein, "subtilisin" means a naturally-occurring subtilisin or a recombinant subtilisin. A series of naturally-occurring subtilisins is known to be produced and often secreted by various microbial species. Amino acid sequences of the members of this series are not entirely homologous. However, the subtilisins in this series exhibit the same or similar type of proteolytic activity. This class of serine proteases shares a common amino acid sequence defining a catalytic triad which distinguishes them from the chymotrypsin related class of serine proteases. The subtilisins and chymotrypsin related serine proteases both have a catalytic triad comprising aspartate, histidine and serine. In the subtilisin related proteases the relative order of these amino acids, reading from the amino to carboxy terminus, is aspartate-histidine-serine. In the chymotrypsin related proteases, the relative order, however, is histidine-aspartate-serine. Thus, subtilisin herein refers to a serine protease having the catalytic triad of subtilisin related proteases. Examples include but are not limited to the subtilisins identified in Fig. 3 herein. Generally and for purposes of the present invention, numbering of the amino acids in proteases corresponds to the numbers assigned to the mature *Bacillus amyloliquefaciens* subtilisin sequence presented in Fig. 1.

"Recombinant subtilisin" or "recombinant protease" refer to a subtilisin or protease in which the DNA sequence encoding the subtilisin or protease is modified to produce a variant (or mutant) DNA sequence which encodes the substitution, deletion or insertion of

on or more amino acids in the naturally-occurring amino acid sequence. Suitable methods to produce such modification, and which may be combined with those disclosed herein, include those disclosed in US Patent 4,760,025 (RE 34,606), US Patent 5,204,015 and US Patent 5,185,258.

5 "Non-human subtilisins" and the DNA encoding them may be obtained from many procaryotic and eucaryotic organisms. Suitable examples of procaryotic organisms include gram negative organisms such as *E. coli* or *Pseudomonas* and gram positive bacteria such as *Micrococcus* or *Bacillus*. Examples of eucaryotic organisms from which subtilisin and their genes may be obtained include yeast such as *Saccharomyces*  
10 *cerevisiae*, fungi such as *Aspergillus* sp.

"Human subtilisin" means the protein represented by the sequence in Fig. 6, derivatives thereof or modifications thereof which retain the essential ability to hydrolyze peptide bonds.

A "protease variant" has an amino acid sequence which is derived from the amino  
15 acid sequence of a "precursor protease". The precursor proteases include naturally-occurring proteases and recombinant proteases. The amino acid sequence of the protease variant is "derived" from the precursor protease amino acid sequence by the substitution, deletion or insertion of one or more amino acids of the precursor amino acid sequence. Such modification is of the "precursor DNA sequence" which encodes the  
20 amino acid sequence of the precursor protease rather than manipulation of the precursor protease enzyme *per se*. Suitable methods for such manipulation of the precursor DNA sequence include methods disclosed herein, as well as methods known to those skilled in the art (see, for example, EP 0 328299, WO89/06279 and the US patents and applications already referenced herein).

25 These amino acid position numbers used herein refer to those assigned to the mature *Bacillus amyloliquefaciens* subtilisin sequence presented in Fig. 1. The invention, however, is not limited to the mutation of this particular subtilisin but extends to precursor proteases containing amino acid residues at positions which are "equivalent" to the particular identified residues in *Bacillus amyloliquefaciens* subtilisin. In a preferred  
30 embodiment of the present invention, the precursor protease is *Bacillus lentus* subtilisin and the substitutions, deletions or insertions are made at the equivalent amino acid residue in *B. lentus* corresponding to those listed above.

A residue (amino acid) of a precursor protease is equivalent to a residue of  
35 *Bacillus amyloliquefaciens* subtilisin if it is either homologous (i.e., corresponding in

position in either primary or tertiary structure) or analogous to a specific residue or portion of that residue in *Bacillus amyloliquefaciens* subtilisin (i.e., having the same or similar functional capacity to combine, react, or interact chemically).

In order to establish homology to primary structure, the amino acid sequence of a precursor protease is directly compared to the *Bacillus amyloliquefaciens* subtilisin primary sequence and particularly to a set of residues known to be invariant in subtilisins for which sequence is known. For example, Fig. 2 herein shows the conserved residues as between *B. amyloliquefaciens* subtilisin and *B. lentus* subtilisin. After aligning the conserved residues, allowing for necessary insertions and deletions in order to maintain alignment (i.e., avoiding the elimination of conserved residues through arbitrary deletion and insertion), the residues equivalent to particular amino acids in the primary sequence of *Bacillus amyloliquefaciens* subtilisin are defined. Alignment of conserved residues preferably should conserve 100% of such residues. However, alignment of greater than 75% or as little as 50% of conserved residues is also adequate to define equivalent residues. Conservation of the catalytic triad, Asp32/His64/Ser221 should be maintained.

For example, in Fig. 6 the amino acid sequence of subtilisin from *Bacillus amyloliquefaciens*, *Bacillus subtilis*, *Bacillus licheniformis* (carlsbergensis) and *Bacillus lentus* are aligned to provide the maximum amount of homology between amino acid sequences. A comparison of these sequences shows that there are a number of conserved residues contained in each sequence. These conserved residues (as between BPN' and *B. lentus*) are identified in Fig. 2.

These conserved residues, thus, may be used to define the corresponding equivalent amino acid residues of *Bacillus amyloliquefaciens* subtilisin in other subtilisins such as subtilisin from *Bacillus lentus* (PCT Publication No. W089/06279 published July 13, 1989), the preferred protease precursor enzyme herein, or the subtilisin referred to as PB92 (EP 0 328 299), which is highly homologous to the preferred *Bacillus lentus* subtilisin. The amino acid sequences of certain of these subtilisins are aligned in Figs. 3A and 3B with the sequence of *Bacillus amyloliquefaciens* subtilisin to produce the maximum homology of conserved residues. As can be seen, there are a number of deletions in the sequence of *Bacillus lentus* as compared to *Bacillus amyloliquefaciens* subtilisin. Thus, for example, the equivalent amino acid for Val165 in *Bacillus amyloliquefaciens* subtilisin in the other subtilisins is isoleucine for *B. lentus* and *B. licheniformis*.

Thus, for example, the amino acid at position +170 is lysine (K) in both *B. amyloliquefaciens* and *B. licheniformis* subtilisins and arginine (R) in Savinas. In the protease variants of the invention, however, the amino acid equivalent to +170 in *Bacillus*

*amyoliquefaciens* subtilisin is substituted with aspartic acid (D). The abbreviations and one letter codes for all amino acids in the present invention conform to the Patent User Manual (GenBank, Mountain View, CA) 1990, p.101.

"Equivalent residues" may also be defined by determining homology at the level of tertiary structure for a precursor protease whose tertiary structure has been determined by x-ray crystallography. Equivalent residues are defined as those for which the atomic coordinates of two or more of the main chain atoms of a particular amino acid residue of the precursor protease and *Bacillus amyoliquefaciens* subtilisin (N on N, CA on CA, C on C and O on O) are within 0.13nm and preferably 0.1nm after alignment. Alignment is achieved after the best model has been oriented and positioned to give the maximum overlap of atomic coordinates of non-hydrogen protein atoms of the protease in question to the *Bacillus amyoliquefaciens* subtilisin. The best model is the crystallographic model giving the lowest R factor for experimental diffraction data at the highest resolution available.

$$R\ factor = \frac{\sum_h |F_o(h)| - |F_c(h)|}{\sum_h |F_o(h)|}$$

Equivalent residues which are functionally analogous to a specific residue of *Bacillus amyoliquefaciens* subtilisin are defined as those amino acids of the precursor protease which may adopt a conformation such that they either alter, modify or contribute to protein structure, substrate binding or catalysis in a manner defined and attributed to a specific residue of the *Bacillus amyoliquefaciens* subtilisin. Further, they are those residues of the precursor protease (for which a tertiary structure has been obtained by x-ray crystallography) which occupy an analogous position to the extent that, although the main chain atoms of the given residue may not satisfy the criteria of equivalence on the basis of occupying a homologous position, the atomic coordinates of at least two of the side chain atoms of the residue lie within 0.13nm of the corresponding side chain atoms of *Bacillus amyoliquefaciens* subtilisin. The coordinates of the three dimensional structure of *Bacillus amyoliquefaciens* subtilisin are set forth in EPO Publication No. 0 251 446 (equivalent to US Patent 5,182,204, the disclosure of which is incorporated herein by reference) and can be used as outlined above to determine equivalent residues on the level of tertiary structure.

Some of the residues identified for substitution, insertion or deletion are conserved residues whereas others are not. In the case of residues which are not conserved, the replacement of one or more amino acids is limited to substitutions which produce a variant which has an amino acid sequence that does not correspond to one found in nature. In the case of conserved residues, such replacements should not result in a naturally-occurring sequence. The protease variants of the present invention include the mature forms of protease variants, as well as the pro- and prepro-forms of such protease variants. The prepro-forms are the preferred construction since this facilitates the expression, secretion and maturation of the protease variants.

"Prosequence" refers to a sequence of amino acids bound to the N-terminal portion of the mature form of a protease which when removed results in the appearance of the "mature" form of the protease. Many proteolytic enzymes are found in nature as translational proenzyme products and, in the absence of post-translational processing, are expressed in this fashion. A preferred prosequence for producing protease variants is the putative prosequence of *Bacillus amyloliquefaciens* subtilisin, although other protease prosequences may be used.

A "signal sequence" or "presequence" refers to any sequence of amino acids bound to the N-terminal portion of a protease or to the N-terminal portion of a proprotease which may participate in the secretion of the mature or pro forms of the protease. This definition of signal sequence is a functional one, meant to include all those amino acid sequences encoded by the N-terminal portion of the protease gene which participate in the effectuation of the secretion of protease under native conditions. The present invention utilizes such sequences to effect the secretion of the protease variants as defined herein. One possible signal sequence comprises the first seven amino acid residues of the signal sequence from *Bacillus subtilis* subtilisin fused to the remainder of the signal sequence of the subtilisin from *Bacillus lentus* (ATCC 21536).

A "prepro" form of a protease variant consists of the mature form of the protease having a prosequence operably linked to the amino terminus of the protease and a "pre" or "signal" sequence operably linked to the amino terminus of the prosequence.

"Expression vector" refers to a DNA construct containing a DNA sequence which is operably linked to a suitable control sequence capable of effecting the expression of said DNA in a suitable host. Such control sequences include a promoter to effect transcription, an optional operator sequence to control such transcription, a sequence encoding suitable mRNA ribosome binding sites and sequences which control termination of transcription and translation. The vector may be a plasmid, a phage particle, or simply a potential



genomic insert. Once transformed into a suitable host, the vector may replicate and function independently of the host genome, or may, in some instances, integrate into the genome itself. In the present specification, "plasmid" and "vector" are sometimes used interchangeably as the plasmid is the most commonly used form of vector at present.

5 However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which are, or become, known in the art.

The "host cells" used in the present invention generally are procaryotic or eucaryotic hosts which preferably have been manipulated by the methods disclosed in US Patent 4,760,025 (RE 34,606) to render them incapable of secreting enzymatically active endoprotease. A preferred host cell for expressing protease is the *Bacillus* strain BG2036  
10 which is deficient in enzymatically active neutral protease and alkaline protease (subtilisin). The construction of strain BG2036 is described in detail in US Patent 5,264,366. Other host cells for expressing protease include *Bacillus subtilis* 1168 (also described in US Patent 4,760,025 (RE 34,606) and US Patent 5,264,366, the disclosure of  
15 which are incorporated herein by reference), as well as any suitable *Bacillus* strain such as *B. licheniformis*, *B. lentus*, etc.

Host cells are transformed or transfected with vectors constructed using recombinant DNA techniques. Such transformed host cells are capable of either replicating vectors encoding the protease variants or expressing the desired protease  
20 variant. In the case of vectors which encode the pre- or prepro-form of the protease variant, such variants, when expressed, are typically secreted from the host cell into the host cell medium.

"Operably linked," when describing the relationship between two DNA regions, simply means that they are functionally related to each other. For example, a  
25 presequence is operably linked to a peptide if it functions as a signal sequence, participating in the secretion of the mature form of the protein most probably involving cleavage of the signal sequence. A promoter is operably linked to a coding sequence if it controls the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to permit translation.

30 The genes encoding the naturally-occurring precursor protease may be obtained in accord with the general methods known to those skilled in the art. The methods generally comprise synthesizing labeled probes having putative sequences encoding regions of the protease of interest, preparing genomic libraries from organisms expressing the protease, and screening the libraries for the gene of interest by hybridization to the probes.

35 Positively hybridizing clones are then mapped and sequenced.

The cloned protease is then used to transform a host cell in order to express the protease. The protease gene is then ligated into a high copy number plasmid. This plasmid replicates in hosts in the sense that it contains the well-known elements necessary for plasmid replication: a promoter operably linked to the gene in question (which may be supplied as the gene's own homologous promoter if it is recognized, i.e., transcribed, by the host), a transcription termination and polyadenylation region (necessary for stability of the mRNA transcribed by the host from the protease gene in certain eucaryotic host cells) which is exogenous or is supplied by the endogenous terminator region of the protease gene and, desirably, a selection gene such as an antibiotic resistance gene that enables continuous cultural maintenance of plasmid-infected host cells by growth in antibiotic-containing media. High copy number plasmids also contain an origin of replication for the host, thereby enabling large numbers of plasmids to be generated in the cytoplasm without chromosomal limitations. However, it is within the scope herein to integrate multiple copies of the protease gene into host genome. This is facilitated by procaryotic and eucaryotic organisms which are particularly susceptible to homologous recombination.

In one embodiment, the gene can be a natural gene such as that from *B. lentus* or *B. amyloliquefaciens*. Alternatively, a synthetic gene encoding a naturally-occurring or mutant precursor protease may be produced. In such an approach, the DNA and/or amino acid sequence of the precursor protease is determined. Multiple, overlapping synthetic single-stranded DNA fragments are thereafter synthesized, which upon hybridization and ligation produce a synthetic DNA encoding the precursor protease. An example of synthetic gene construction is set forth in Example 3 of US Patent 5,204,015, the disclosure of which is incorporated herein by reference.

Once the naturally-occurring or synthetic precursor protease gene has been cloned, a number of modifications are undertaken to enhance the use of the gene beyond synthesis of the naturally-occurring precursor protease. Such modifications include the production of recombinant proteases as disclosed in US Patent 4,760,025 (RE 34,606) and EPO Publication No. 0 251 446 and the production of protease variants described herein.

The following cassette mutagenesis method may be used to facilitate the construction of the protease variants of the present invention, although other methods may be used. First, the naturally-occurring gene encoding the protease is obtained and sequenced in whole or in part. Then the sequence is scanned for a point at which it is desired to make a mutation (deletion, insertion or substitution) of one or more amino acids

in the encoded enzyme. The sequences flanking this point are evaluated for the presence of restriction sites for replacing a short segment of the gene with an oligonucleotide pool which when expressed will encode various mutants. Such restriction sites are preferably unique sites within the protease gene so as to facilitate the replacement of the gene segment. However, any convenient restriction site which is not overly redundant in the protease gene may be used, provided the gene fragments generated by restriction digestion can be reassembled in proper sequence. If restriction sites are not present at locations within a convenient distance from the selected point (from 10 to 15 nucleotides), such sites are generated by substituting nucleotides in the gene in such a fashion that neither the reading frame nor the amino acids encoded are changed in the final construction. Mutation of the gene in order to change its sequence to conform to the desired sequence is accomplished by M13 primer extension in accord with generally known methods. The task of locating suitable flanking regions and evaluating the needed changes to arrive at two convenient restriction site sequences is made routine by the redundancy of the genetic code, a restriction enzyme map of the gene and the large number of different restriction enzymes. Note that if a convenient flanking restriction site is available, the above method need be used only in connection with the flanking region which does not contain a site.

Once the naturally-occurring DNA or synthetic DNA is cloned, the restriction sites flanking the positions to be mutated are digested with the cognate restriction enzymes and a plurality of end termini-complementary oligonucleotide cassettes are ligated into the gene. The mutagenesis is simplified by this method because all of the oligonucleotides can be synthesized so as to have the same restriction sites, and no synthetic linkers are necessary to create the restriction sites.

In one aspect of the invention, the objective is to secure a variant protease having altered allergenic potential as compared to the precursor protease, since decreasing such potential enables safer use of the enzyme. While the instant invention is useful to lower allergenic potential, the mutations specified herein may be utilized in combination with mutations known in the art to result altered thermal stability and/or altered substrate specificity, modified activity or altered alkaline stability as compared to the precursor.

Thus, in combination with the mutations of the present invention, substitutions at positions corresponding to N76D/S103A/V104I/G159D optionally in combination with one or more substitutions selected from the group consisting of positions corresponding to V68A, T213R, A232V, Q236H, Q245R, and T260A of *Bacillus amyloliquefaciens* subtilisin may be used, in addition to decreasing the allergenic potential of the variant protease of

the invention, to modulate overall stability and/or proteolytic activity of the enzyme. Similarly, the substitutions provided herein may be combined with mutation at the Asparagine (N) in *Bacillus lentus* subtilisin at equivalent position +76 to Aspartate (D) in combination with the mutations S103A/V104I/G159D and optionally in combination with one or more substitutions selected from the group consisting of positions corresponding to V68A, T213R, A232V, Q236H, Q245R, and T260A of *Bacillus amyloliquefaciens* subtilisin, to produce enhanced stability and/or enhanced activity of the resulting mutant enzyme.

Based on the screening results obtained with the variant proteases, the noted mutations in *Bacillus amyloliquefaciens* subtilisin are important to the proteolytic activity, performance and/or stability of these enzymes and the cleaning or wash performance of such variant enzymes.

Many of the protease variants of the invention are useful in formulating various detergent compositions. A number of known compounds are suitable surfactants useful in compositions comprising the protease mutants of the invention. These include nonionic, anionic, cationic, anionic or zwitterionic detergents, as disclosed in US 4,404,128 to Barry J. Anderson and US 4,261,868 to Jiri Flora, et al. A suitable detergent formulation is that described in Example 7 of US Patent 5,204,015 (previously incorporated by reference). The art is familiar with the different formulations which can be used as cleaning compositions. In addition to typical cleaning compositions, it is readily understood that the protease variants of the present invention may be used for any purpose that native or wild-type proteases are used. Thus, these variants can be used, for example, in personal care items such as face lotions and cosmetics, in bar or liquid soap applications, dishcare formulations, contact lens cleaning solutions or products, peptide hydrolysis, waste treatment, textile applications, as fusion-cleavage enzymes in protein production, etc. The variants of the present invention may comprise enhanced performance in a detergent composition (as compared to the precursor). As used herein, enhanced performance in a detergent is defined as increasing cleaning of certain enzyme sensitive stains such as grass or blood, as determined by usual evaluation after a standard wash cycle.

Proteases of the invention can be formulated into known powdered and liquid detergents having pH between 6.5 and 12.0 at levels of about .01 to about 5% (preferably .1% to .5%) by weight. These detergent cleaning compositions can also include other enzymes such as known proteases, amylases, cellulases, lipases or endoglycosidases, as well as builders and stabilizers.

The addition of proteases of the invention to conventional cleaning compositions does not create any special use limitation. In other words, any temperature and pH

suitable for the detergent is also suitable for the present compositions as long as the pH is within the above range, and the temperature is below the described protease's denaturing temperature. In addition, proteases of the invention can be used in a cleaning composition without detergents, again either alone or in combination with builders and stabilizers.

The variant proteases of the present invention can be included in animal feed such as part of animal feed additives as described in, for example, US 5,612,055; US 5,314,692; and US 5,147,642.

One aspect of the invention is a composition for the treatment of a textile that includes variant proteases of the present invention. The composition can be used to treat for example silk or wool as described in publications such as RD 216,034; EP 134,267; US 4,533,359; and EP 344,259.

The following is presented by way of example and is not to be construed as a limitation to the scope of the claims.

All publications and patents referenced herein are hereby incorporated by reference in their entirety.

## **EXAMPLES**

### **Example 1**

#### **Assay for the Identification of Peptide T-Cell Epitopes Using Naïve Human T-Cells**

Fresh human peripheral blood cells were collected from "naïve" humans, i.e., persons not known to be exposed to or sensitized to *Bacillus lentus* protease, for determination of antigenic epitopes in protease from *Bacillus lentus* and human subtilisin. Naïve humans is intended to mean that the individual is not known to have been exposed to or developed a reaction to protease in the past. Peripheral mononuclear blood cells (stored at room temperature, no older than 24 hours) were prepared for use as follows: Approximately 30 mls of a solution of buffy coat preparation from one unit of whole blood was brought to 50 ml with Dulbecco's phosphate buffered solution (DPBS) and split into two tubes. The samples were underlaid with 12.5 ml of room temperature lymphoprep density separation media (Nycomed density 1.077 g/ml). The tubes were centrifuged for thirty minutes at 600G. The interface of the two phases was collected, pooled and washed in DPBS. The cell density of the resultant solution was measured by hemocytometer. Viability was measured by trypan blue exclusion.

From the resulting solution, a differentiated dendritic cell culture was prepared from the peripheral blood mononuclear cell sample having a density of  $10^8$  cells per 75 ml culture flask in a solution as follows:

(1) 50 ml of serum free AIM V media (Gibco) was supplemented with a 1:100 dilution beta-mercaptoethanol (Gibco). The flasks were laid flat for two hours at 37°C in 5% CO<sub>2</sub> to allow adherence of monocytes to the flask wall.

(2) Differentiation of the monocyte cells to dendritic cells was as follows: nonadherent cells were removed and the resultant adherent cells (monocytes) combined with 30 ml of AIM V, 800 units/ml of GM-CSF (Endogen) and 500 units/ml of IL-4 (Endogen); the resulting mixture was cultured for 5 days under conditions at 37°C in 5% CO<sub>2</sub>. After five days, the cytokine TNF( $\alpha$ ) (Endogen) was added to 0.2 units/ml, and the cytokine IL-1 $\alpha$  (Endogen) was added to a final concentration of 50 units/ml and the mixture incubated at 37°C in 5% CO<sub>2</sub> for two more days.

(3) On the seventh day, Mitomycin C was added to a concentration of 50 microgram/ml was added to stop growth of the now differentiated dendritic cell culture. The solution was incubated for 60 minutes at 37°C in 5% CO<sub>2</sub>. Dendritic cells were collected by gently scraping the adherent cells off the bottom of the flask with a cell scraper. Adherent and non-adherent cells were then centrifuged at 600G for 5 minutes, washed in DPBS and counted.

(4) The prepared dendritic cells were placed into a 96 well round bottom array at  $2 \times 10^4$ /well in 100 microliter total volume.

CD4+ T cells were prepared from frozen aliquots of the peripheral blood cell samples used to prepare the dendritic cells using the human CD4+ Collect Kit (Biotex) as per the manufacturers instructions with the following modifications: the aliquots were thawed and washed such that approximately  $10^8$  cells will be applied per Collect column; the cells were resuspended in 4 ml DPBS and 1 ml of the Cell reagent from the Collect Kit, the solution maintained at room temperature for 20 minutes. The resultant solution was centrifuged for five minutes at 600G at room temperature and the pellet resuspended in 2 ml of DPBS and applied to the Collect columns. The effluent from the columns was collected in 2% human serum in DPBS. The resultant CD4+ cell solution was centrifuged, resuspended in AIMV media and the density counted.

The CD4+ T-cell suspension was resuspended to a count of  $2 \times 10^6$ /ml in AIM V media to facilitate efficient manipulation of the 96 well plate.

Peptide antigen is prepared from a 1M stock solution in DMSO by dilution in AIM V media at a 1:10 ratio. 10 microliters of the stock solution is placed in each well of the 96 well plate containing the differentiated dendritic cells. 100 microliter of the diluted CD4+ T-cell solution as prepared above is further added to each well. Useful controls include diluted DMSO blanks, and tetanus toxoid positive controls.

The final concentrations in each well, at 210 microliter total volume are as follows:

$2 \times 10^5$  CD4+

$2 \times 10^4$  dendritic cells (R:S of 10:1)

5 mM/ $10^4$  peptide

## Example 2

### Identification of T-Cell Epitopes in Protease from *Bacillus lentus* and Human subtilisin

Peptides for use in the assay described in Example 1 were prepared based on the *Bacillus lentus* and human subtilisin amino acid sequence. Peptide antigens were designed as follows. From the full length amino acid sequence of either human subtilisin or *Bacillus lentus* protease provided in Figure 1, 15mers were synthetically prepared, each 15mer overlapping with the previous and the subsequent 15mer except for three residues.

Peptides used correspond to amino acid residue strings in *Bacillus lentus* as provided in Figure 7, and peptides correspond to amino acid residues in human subtilisin as provided in Figure 8. The key for the coded results is provided in Figure 10. All tests were performed at least in duplicate. All tests reported displayed robust positive control responses to the antigen tetanus toxoid. Responses were averaged within each experiment, then normalized to the baseline response. A positive event was recorded if the response was at least 3 times the baseline response.

The immunogenic response (i.e., T-cell proliferation) to the prepared peptides from human subtilisin and *Bacillus lentus* was tallied and is provided in Figures 4 and 5, respectively. T-cell proliferation was measured by the incorporated tritium method. The results shown in Figures 4 and 5 as a comparison of the immunogenic additive response in 10 individuals (figure 4) and 16 individuals (figure 5) to the various peptides. Response is indicated as the added response wherein 1.0 equals a baseline response for each sample. Thus, in Figure 4, a reading of 10.0 or less is the baseline response and in Figure 5 a reading of 16.0 or less the baseline response.

As indicated in Figures 4 and 5, the immunogenic response of the naïve blood samples from unsensitized individuals showed a marked allergenic response at the peptide fragment from *Bacillus lentus* corresponding to residues 170-173 of *Bacillus amyloliquefaciens* protease. As expected, the corresponding fragment in human subtilisin evokes merely baseline response.

5



**WE CLAIM:**

1. A method for reducing the allergenicity of a non-human protein is provided wherein an epitope is identified and replaced with an analogous region within a human subtilisin.

5           2. The method according to claim 1, wherein said non-human protein comprises a protease.

3. The method according to claim 2, wherein said epitope replaced is a T-cell epitope.

10           4. A method for producing the protein of the invention having reduced allergenicity comprising, preparing a mutant protein by modifying a DNA encoding a precursor protein so that the modified DNA encodes the mutant protein of the invention, wherein an epitope is replaced with an analogous region from human subtilisin.

5. A mutant protein having reduced allergenicity produced according to the method of claim 4.

15           6. DNA encoding the mutant protein of claim 5.

7. An expression vector containing a DNA sequence according to claim 6.

8. Host cells transformed with the vectors of claim 7.

20           9. A method of producing a protein having reduced allergenicity wherein the host cells of claim 8 are cultivated under conditions which result in expression of the mutant protease.

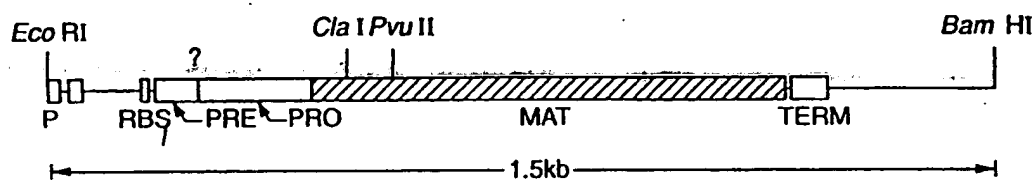
10. A pharmaceutical composition comprising a human subtilisin.

11. A DNA encoding a human subtilisin.

12. A vector comprising the DNA of claim 11.

13. A host cell comprising the vector of claim 12.

25           14. A human subtilisin produced by the host cell of claim 12.

**FIG. 1A**

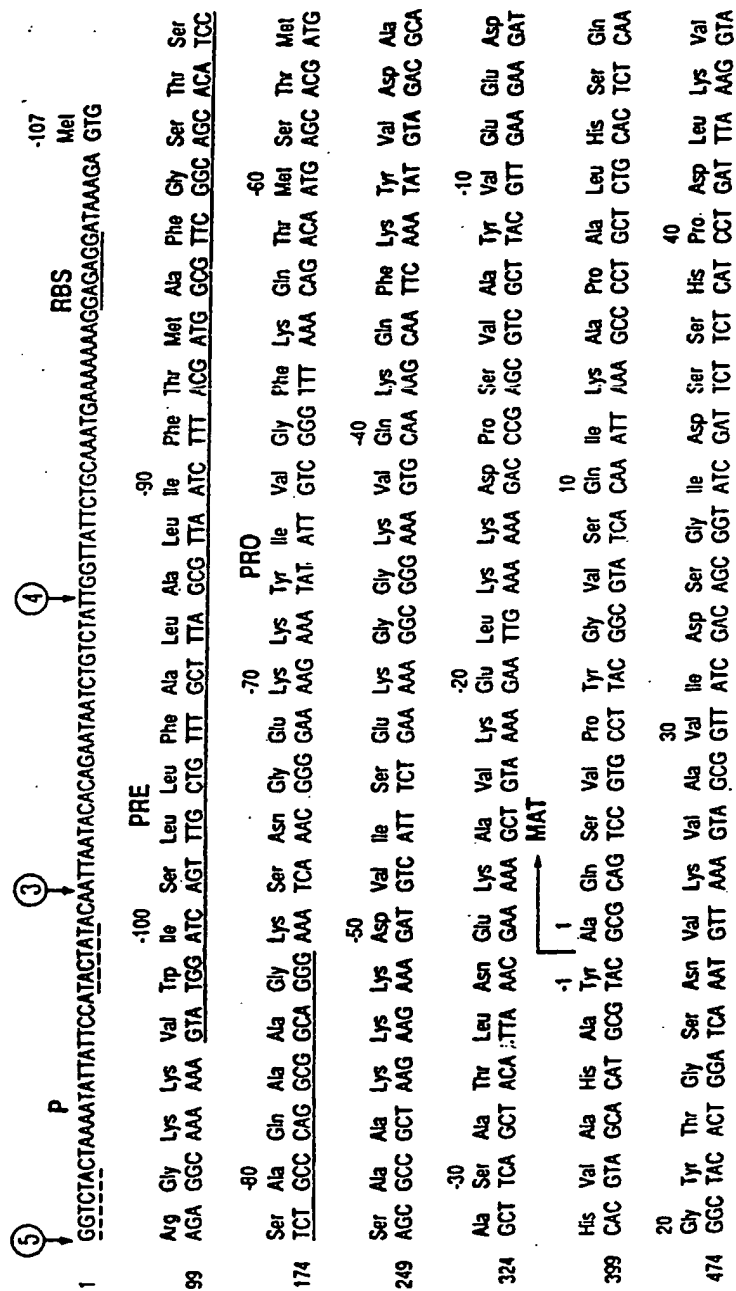


FIG. 1B - 1

3/13

50      Pro   Asn   Asp   His   Gly   Thr   His   Val   Ala  
 Ala   Gly   Gly   GGA   GCC   AGC   ATG   GTT   CCT   TCT   TCT   TCT   TCT   CAC   GGA   ACT   CAC   GTT   GCC  
 549   GCA   GGC   GGA   GGC   GGC   GGC   GGC   GGC   GGC   GGC   GGC   GGC   GGC   GGC   GGC   GGC   GGC   GGC   GGC  
 60      Phe   Gln   Asp   Asn   Asn   Ser   His   Gly   Thr   His   Val   Ala  
 Thr   Ser   Gln   Thr   Thr   Thr   Thr   Thr   Thr   Thr   Thr   Thr   Thr   Thr   Thr   Thr   Thr   Thr  
 70      Gly   Thr   Val   Ala   Ala   Leu   Asn   Asn   Ser   Ile   Ile   Gly   Val   Ala   Pro   Ser   Ala   Ser   Leu   Tyr   Ala   Val   Lys  
 824   GGC   ACA   GTT   GCG   GGT   CTT   AAT   AAC   TCA   ATC   GGT   GTA   TTA   GGC   GTT   GCG   CCA   AGC   GCA   TCA   CTT   TAC   GCT   GTA   AAA  
 80      Ser   Ala   90  
 100      Asp   Ala   100  
 Val   Leu   Gly   Ala   Asp   Gly   Ser   Gly   Gln   Tyr   Ser   Trp   Ile   Ile   Asn   Gly   Ile   Gln   Trp   Ala   Ile   Ala   Asn   Asn   Met  
 699   GTT   CTC   GGT   GGT   GAC   GGT   TCC   TCC   GGC   CAA   TAC   AGC   TGG   ATC   ATT   AAC   GGA   ATC   GAG   TGG   GCG   ATC   GCA   AAC   AAT   ATG  
 110      110  
 120      Asp   Val   Ile   Asn   Met   Ser   Leu   Gly   Gly   Pro   Ser   Gly   Ser   Ala   Ala   Leu   Lys   Ala   Val   Asp   Lys   Ala   Val   Ala  
 774   GAC   GTT   ATT   AAC   ATG   ATG   AGC   CTC   GGC   GGA   CCT   TCT   GGT   TCT   GCT   GCT   TTA   AAA   GCG   GCA   GTT   GAT   AAA   GCC   GTT   GCA  
 130      130  
 140      Ser   Thr   140  
 150      Ser   Thr   150  
 Ser   Gly   Val   Val   Val   Val   Ala   Ala   Gly   Asn   Gln   Gly   Thr   Ser   Gly   Ser   Ser   Thr   Val   Val   Gly   Tyr   Pro   Gly  
 849   TCC   GGC   GTC   GTA   GTC   GTT   GCG   GCA   GCC   GGT   AAC   GAA   GGC   ACT   TCC   GGC   AGC   TCA   AGC   ACA   GTG   GGC   TAC   CCT   GGT  
 160      160  
 170      Lys   Tyr   Pro   Ser   Val   Ile   Ala   Val   Gly   Ala   Val   Asp   Ser   Ser   Asn   Gln   Arg   Ala   Ser   Phe   Ser   Ser   Val   Gly   Pro  
 924   AAA   TAC   CCT   TCT   GTC   ATT   GCA   GTA   GGC   GGT   GTT   GAC   AGC   AGC   AAC   CAA   AGA   GCA   TCT   TTC   TCA   AGC   GTA   GGA   CCT  
 180      180  
 190      Gly   Leu   Asp   Val   Met   Ala   Pro   Gly   Val   Ser   Ile   Gln   Ser   Thr   Leu   Pro   Gly   Asn   Lys   Tyr   Gly   Ala   Tyr   Asn   Gly  
 999   GAG   CTT   GAT   GTC   ATG   GCA   CCT   GGC   GTA   TCT   ATC   CAA   AGC   ACG   CTT   CCT   GGA   AAC   AAA   TAC   TAC   GGG   GCG   TAC   AAC   GGT  
 200      200  
 210      Thr   Ser   Met   Ala   Ser   Pro   His   Val   Val   Ala   Gly   Ala   Ala   Ala   Leu   Ile   Leu   Ser   Lys   His   Pro   Asn   Trp   Thr   Asn   Thr  
 1074   ACG   TCA   ATG   GCA   TCT   CCG   CAC   GGT   GCG   GCA   GCG   GGT   GGT   TTG   ATT   CTT   TCT   AAG   CAC   CCG   AAC   TGG   ACA   AAC   ACT  
 220      220  
 230      Thr   Ser   Met   Ala   Ser   Pro   His   Val   Val   Ala   Gly   Ala   Ala   Ala   Leu   Ile   Leu   Ser   Lys   His   Pro   Asn   Trp   Thr   Asn   Thr  
 240      240

FIG. 1B - 2

4/13

250 Gln  
Gln Val Arg Ser Ser Leu Glu Asn Thr Thr Thr Lys Leu Gly Asp Ser Phe Tyr Tyr Gly Lys Gly Leu Ile Asn  
1149 CAA GTC CGC AGC AGT TTA GAA AAC ACC ACT ACA AAA CTT GGT GAT TCT TTC TAC TAT GGA AAA GGG CTG ATC AAC

270 Val Gln Ala Ala Ala Gln OC  
1224 GTA CAG GCG GCA GCT CAG TAA AACAATAAAACCGGCCTTGGCCCCCGCGGGTTTTTATTTTCTTCCTCCGATGTTCAATCGGCTCC TERM

1316 ATAATCGACGGATGGCTCCCTCTGAAAAATTTAACGAGAACGGGGGTTGACCCGGCTCAGTCCCGTAACGGCCAAAGTCCCTGAACGTCTCAATCGCGG

1416 CTTCGGGTTCCGGTCAGCTCAATGCCGTAACGGTCGGCGCGTTTTCCTGATACCGGAGACGGCATTTCGTAATCGGATC

FIG.\_1B - 3

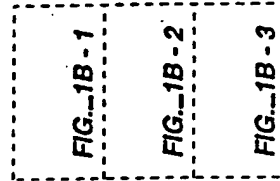


FIG.\_1B

5/13

CONSERVED RESIDUES IN SUBTILISINS FROM  
BACILLUS AMYLOLIQUEFACIENS

```

1           10           20
A Q S V P . G . . . . . A P A . H . . G

21          30          40
. T G S . V K V A V . D . G . . . . H P

41          50          60
D L . . . . G G A S . V P . . . . . Q D

61          70          80
. N . H G T H V A G T . A A L N N S I G

81          90          100
V L G V A P S A . L Y A V K V L G A . G

101         110         120
S G . . S . L . . G . E W A . N . . . .

121         130         140
V . N . S L G . P S . S . . . . . A . .

141         150         160
. . . . . G V . V V A A . G N . G . . . .

161         170         180
. . . . . Y P . . Y . . . . A V G A .

181         190         200
D . . N . . A S P S . . G . . L D . . A

201         210         220
P G V . . Q S T . P G . . Y . . . . N G T

221         230         240
S M A . P H V A G A A A L . . . . K . . . .

241         250         260
W . . . Q . R . . L . N T . . . . L G . . .

261         270
. . Y G . G L . N . . A A . .

```

FIG.\_2

6/13

## COMPARISON OF SUBTILISIN SEQUENCES FROM:

*B. amyloliquefaciens**B. subtilis**B. licheniformis**B. lentus*

01	10	20	30	
A Q S V P Y G V S Q I K A P A L H S Q G Y T G S N V K V A V I D S S S H P P				
A Q S V P Y G I S Q I K A P A L H S Q G Y T G S N V K V A V I D S S S H P P				
A Q T V P Y G I P L I K A D K V Q A Q G F K G A N V K V A V L D T G I Q A S H P P				
A Q S V P W G I S R V Q A P A A H N R G L T G S G V K V A V L D T G I S T * H P				
41	50	60	70	
D L K V A G G A S M V P S E T N P F Q D D N N S H G T H V A G T V A A L N N S I G				
D L N V R G G A S F V P S E T N P Y Q D G S S H G T H V A G T V A A L N N S I G				
D L N V V G G A S F V A G E A Y N * T D G N G H G T H V A G T V A A L D N T T G				
D L N I R G G A S F V P G E * P S T Q D D G N G H G T H V A G T V A A L N N S I G				
81	90	100	110	
V L G V A P S A S L Y A V K V L G A D G S G Q Y S W I I N G I E W A I A N N M D				
V L G V S P S A S L Y A V K V L D S T G S G Q Y S W I I N G I E W A I S N N M D				
V L G V A P S V S L Y A V K V L N S S G S G S Y S G I V S G I E W A T T N G M D				
V L G V A P S A E L Y A V K V L G A S G S G S V S S I A Q G L E W A G N N G M H				
121	130	140	150	
V I N M S L G G P S G S A A L K A A V D K A V A S G V V V A A A A G N E G T S G				
V I N M S L G G P T G S T A L K T V V D K A V S S G I V V A A A A A G N E G S S G				
V I N M S L G G A S G S T A M K Q A V D N A Y A R G V V V V A A A A G N S G N S G				
V A N L S L G S P S P S A T L E Q A V N S A T S R G V L V V A A S G N S G A G S				

FIG.-3A

7/13

161 SS STVGYPPGKYPPSPVIAVGAVDSSNQRA SFS VGP ELDVMA  
170 SS STVGYPPAKYPPSTIAVGAVDSSNQRA SFS AGSELDVMA  
180 SS STVGYPPAKYPPSTIAVGAVDSSNQRA SFS V GAELEVMA  
190 SS STVGYPPAKYPPSTIAVGAVDSSNQRA SFS YGAGLDI V A  
201 PGVSIQSTLPPGNNKYGAYNGTSMASPHVAGAAALILSKHHPN  
210 PGVSIQSTLPPGCTYGCAYNGTSMATPHVAGAAALILSKHHPN  
220 PGAGGVYSTYPPGSTYASLNGTSMATPHVAGAAALVKKKPPS  
230 PGVNVQSTYPPGSTYASLNGTSMATPHVAGAAALVKKKPPS  
241 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
250 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
260 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
270 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
280 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
290 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
300 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
310 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
320 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
330 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
340 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
350 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
360 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
370 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
380 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
390 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
400 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
410 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
420 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
430 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
440 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
450 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
460 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
470 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
480 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
490 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ  
500 WTNATQVRRSSLEENNTTTLGGDSTFY YGKGLIN VQAAAQ

FIG.\_3B

FIG.\_3

FIG.\_3A

FIG.\_3B



8/13

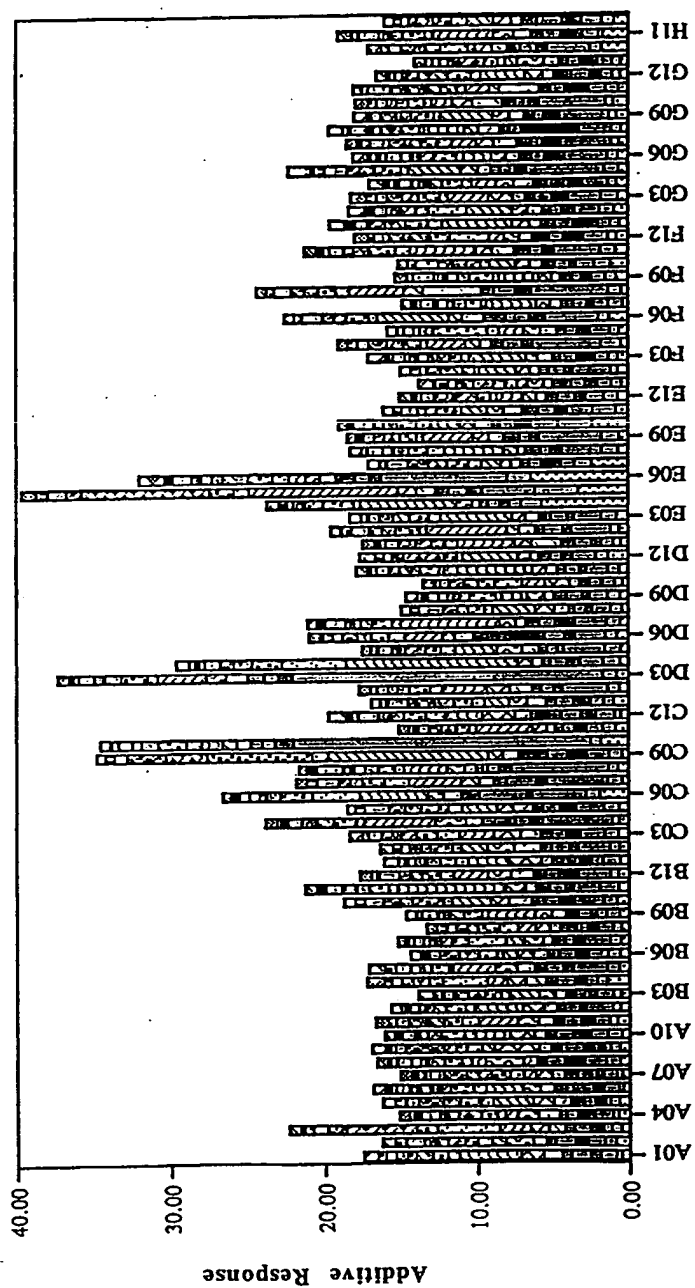


FIG. 4

9/13

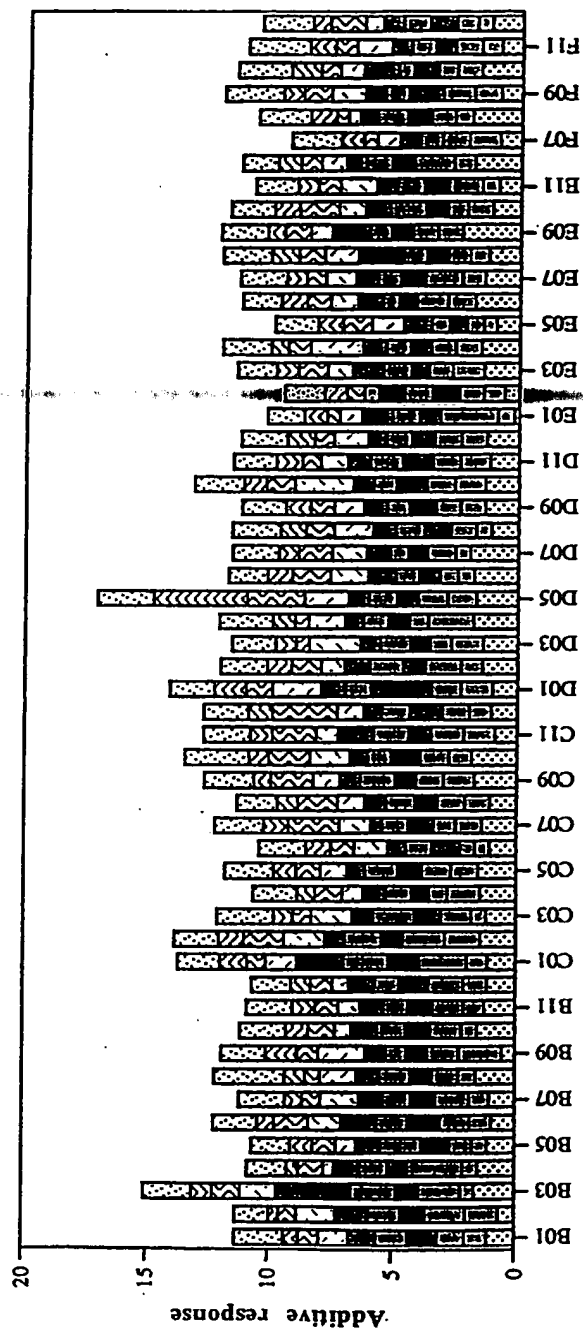


FIG. 5

10/13

MKLVNIWLLLLLVLLCGKKHLGDRLEKKSFEKAPCPGCSHLTLKVEFSSTVVEYEYIVAFNGYFT  
AKARNSFISSALKSSEVDNWRIIPRNNPSSDYPSEFEVIQIKEKQKAGLLTLEDHPNIKRVTPQR  
KVFRSLKYAESDPTVPCNETRWSQKWQSSRPLRRASLSLGS GFWHATGRHSSRLLRAIPRQVAQ  
TLQADVLWQMGYTGANVRVAVFDTGLSEKHPHFKNVKERTNWTNERTLDDGLGHGTFVAGVIASM  
RECQGFAPDAELHIFRVFTNNQVSYTSWFLDAFNAILKKIDVLNLSIGGPDFMDHPFVDK VWEL  
TANNVIMVSAIGNDGPLYGTLNPNADQMDVIGVGGIDFEDNIARFSSRGMTTWELPGGYGRMKPD  
IVTYGAGVRGSGVKGGCRALSGTSVASPVVAGAVTLLVSTVQKREL VNPASMKQALIASARRLP  
VNMFEQGHGKLDLLRAYQILNSYKPQASLSPSYIDLTECPYMWPYCSQPIYYGGMPTVVNVITLN  
GMGVTGRIVDKPDWQPYLPQNGDNIEVAFSYSSVLWPWSGYLAISISVTKKAASWEGIAQGHVMI  
TVASPAETESKNGAEQTSTVKLPIKVKIIPTPPRSKRVLWDQYHNLRYPPGYFPRDNLRMKNDPL  
DWNGDHIHTNFRDMYQHLRSMGYFVEVLGAPFTCFDASQYGTLLMVDSEEEYFPEEIAKLRRDVD  
NGLSLVIFSDWYNTSVMRKVKFYDENTRQWWMPDTGGANIPALNELLSVWNMGFSDGLYEGETL  
ANHDMYYASGCSIAKFPEDGVVITQTFKDQGLEVLKQETAVVENVPILGLYQIPAEGGGRIVLYG  
DSNCLDDSHRQKDCFWLLDALLQYTSYGVTPPSLSHSGNRQRPPSGAGSVTPERMEGNHLHRYSK  
VLEAHLGDPKPRPLPACPRLSWAKPQPLNETAPSNLWKHQKLLSIDLDKVVLPNFRSNRPQVRPL  
SPGESGAWDIPGGIMPGRYNQEVGQTI PVFAFLGAMVVLAFV VQINKAKSRPKRRKPRVKRPQL  
MQQVHPPKTPSV

FIG. 6

1	A12	IKDFHVYFRESRDAG	49	E12	SATSRGVLVVAASGN
2	A11	LEQAVNSATSRGVLV	50	E11	SRGVLVVAASGNSGA
3	A10	AQSVPWGISRVQAPA	51	E10	VLVVAASGNSGAGSI
4	A9	VPWGISRVQAPAAHN	52	E9	VAASGNSGAGSISYP
5	A8	GISRVQAPAAHNRGL	53	E8	SGNSGAGSISYPARY
6	A7	RVQAPAAHNRGLTGS	54	E7	SGAGSISYPARYANA
7	A6	APAAHNRGLTGSGVK	55	E6	GSISYPARYANAMAV
8	A5	AHNRGLTGSGVKVAV	56	E5	SYPARYANAMAVGAT
9	A4	RGLTGSGVKVAVLDT	57	E4	ARYANAMAVGATDQN
10	A3	TGSGVKVAVLDTGIS	58	E3	ANAMAVGATDQNNNR
11	A2	GVKVAVLDTGISTHP	59	E2	MAVGATDQNNNRASF
12	A1	VAVLDTGISTHPDLN	60	E1	GATDQNNNRASFQY
13	B12	LDTGISTHPDLNIRG	61	F12	DQNNNRASFQYGAG
14	B11	GISTHPDLNIRGGAS	62	F11	NNRASFSQYGAGLDI
15	B10	THPDLNIRGGASFVP	63	F10	ASFSQYGAGLDIVAP
16	B9	DLNIRGGASFVPGEF	64	F9	SQYGAGLDIVAPGVN
17	B8	IRGGASFVPGEFSTQ	65	F8	GAGLDIVAPGVNVQS
18	B7	GASFVPGEFSTQDGN	66	F7	LDIVAPGVNVQSTYP
19	B6	FVPGEFSTQDGNHGH	67	F6	VAPGVNVQSTYPGST
20	B5	GEPSTQDGNHGHGTHV	68	F5	GVNVQSTYPGSTYAS
21	B4	STQDGNHGHGTHVAGT	69	F4	VQSTYPGSTYASLNG
22	B3	DGNHGHGTHVAGTIAA	70	F3	TYPGSTYASLNGTSM
23	B2	GHGTHVAGTIAALNN	71	F2	GSTYASLNGTSMATP
24	B1	THVAGTIAALNNSIG	72	F1	YASLNGTSMATPHVA
25	C12	AGTIAALNNSIGVLG	73	G12	LNGTSMATPHVAGAA
26	C11	IAALNNSIGVLGVAP	74	G11	TSMATPHVAGAAALV
27	C10	LNNSIGVLGVAPSAE	75	G10	ATPHVAGAAALVKQK
28	C9	SIGVLGVAPSAELYA	76	G9	HVAGAAALVKQKNPS
29	C8	VLGVAPSAELYAVKV	77	G8	GAAALVKQKNPSWSN
30	C7	VAPSAELYAVKVLGA	78	G7	ALVKQKNPSWSNVQI
31	C6	SAELYAVKVLGASGS	79	G6	KQKNPSWSNVQIRNH
32	C5	LYAVKVLGASGSGSV	80	G5	NPSWSNVQIRNHLKN
33	C4	VKVLGASGSGSVSSI	81	G4	WSNVQIRNHLKNTAT
34	C3	LGASGSGSVSSIAQG	82	G3	VQIRNHLKNTATSLG
35	C2	SGSGSVSSIAQGLEW	83	G2	RNHLKNTATSLGSTN
36	C1	GSVSSIAQGLEWAGN	84	G1	LKNTATSLGSTNLYG
37	D12	SSIAQGLEWAGNNGM	85	H12	TATSLGSTNLYGSGL
38	D11	AQGLEWAGNNGMHVA	86	H11	SLGSTNLYGSGLVNA
39	D10	LEWAGNNGMHVANLS	87	H10	STNLYGSGLVNAEAA
40	D9	AGNNGMHVANLSLGS	88	H9	NLYGSGLVNAEAATR
41	D8	NGMHVANLSLGSPSP			
42	D7	HVANLSLGSPSPSAT			
43	D6	NLSLGSPSPSATLEQ			
44	D5	LGSPSPSATLEQAVN			
45	D4	PSPSATLEQAVNSAT			
46	D3	SATLEQAVNSATSRG			
47	D2	LEQAVNSATSRGVLV			
48	D1	AVNSATSRGVLVVA			

FIG. 7

1	A12	IKDFHVYFRESRDAG	49	E12	KKIDVLNLSIGGPDF
2	A11	DAELHIFRVFTNNQV	50	E11	DVLNLSIGGPDFMDH
3	A10	PLRRASLSLGSGFWH	51	E10	NLSIGGPDFMDHPFV
4	A9	RASLSLGSGFWHATG	52	E9	IGGPDFMDHPFVDKV
5	A8	LSLGSGFWHATGRHS	53	E8	PDFMDHPFVDKVVWEL
6	A7	GSGFWHATGRHSSRR	54	E7	MDHPFVDKVVWELTAN
7	A6	FWHATGRHSSRLLR	55	E6	PFVDKVVWELTANNVI
8	A5	ATGRHSSRLLRAIP	56	E5	DKVVWELTANNVIMVS
9	A4	RHSSRLLRAIPRQV	57	E4	WELTANNVIMVSAIG
10	A3	SRLLRAIPRQVAQT	58	E3	TANNVIMVSAIGNDG
11	A2	LLRAIPRQVAQTLQA	59	E2	NVIMVSAIGNDGPLY
12	A1	AIPRQVAQTLQADVL	60	E1	MVSAIGNDGPLYGTJ
13	B12	RQVAQTLQADVLWQM	61	F12	AIGNDGPLYGTLNPN
14	B11	AQTLQADVLWQMGYT	62	F11	NDGPLYGTLNPNADQ
15	B10	LQADVLWQMGYTGAN	63	F10	PLYGTLNPNADQMDV
16	B9	DVLWQMGYTGANVRV	64	F9	GTLLNPNADQMDVIGV
17	B8	WQMGYTGANVRVAVF	65	F8	NNPADQMDVIGVGGI
18	B7	GYTGANVRVAVFDTG	66	F7	ADQMDVIGVGGIDFE
19	B6	GANVRVAVFDTGLSE	67	F6	MDVIGVGGIDFEDNI
20	B5	VRVAVFDTGLSEKHP	68	F5	IGVGGIDFEDNIARF
21	B4	AVFDTGLSEKHPHFK	69	F4	GGIDFEDNIARFSSR
22	B3	DTGLSEKHPHFKNVK	70	F3	DFEDNIARFSSRGMT
23	B2	LSEKHPHFKNVKERT	71	F2	DNIARFSSRGMTTWE
24	B1	KHPHFKNVKERTNWT	72	F1	ARFSSRGMTTWELPG
25	C12	HFKNVKERTNWTNER	73	G12	SSRGMTTWELPGGYG
26	C11	NVKERTNWTNERTLD	74	G11	GMTTWELPGGYGRMK
27	C10	ERTNWTNERTLDDGL	75	G10	TWELPGGYGRMKPDI
28	C9	NWTNERTLDDGLGHG	76	G9	LPGGYGRMKPDIVTY
29	C8	NERTLDDGLGHGTFV	77	G8	GYGRMKPDIVTYGAG
30	C7	TLDDGLGHGTFVAGV	78	G7	RMKPDIVTYGAGVRG
31	C6	DGLGHGTFVAGVIAS	79	G6	PDIVTYGAGVRGSGV
32	C5	GHGTFVAGVIASMRE	80	G5	VTYAGVRGSGVKGG
33	C4	TFVAGVIASMRECQG	81	G4	GAGVRGSGVKGGCRA
34	C3	AGVIASMRECQGFAP	82	G3	VRGSGVKGGCRALSG
35	C2	IASMRECQGFAPDAE	83	G2	SGVKGGCRALSGTSV
36	C1	MRECQGFAPDAELHI	84	G1	KGGCRALSGTSVASP
37	D12	CQGFAPDAELHIFRV	85	H12	CRALSGTSVASPVVA
38	D11	FAPDAELHIFRVFTN	86	H11	LSGTSVASPVVAGAV
39	D10	DAELHIFRVFTNNQV	87	H10	TSVASPVVAGAVTLL
40	D9	LHIFRVFTNNQVSYT	88	H9	ASPVVAGAVTLLVST
41	D8	FRVFTNNQVSYTSWF	89	H8	VVAGAVTLLVSTVQK
42	D7	FTNNQVSYTSWFLDA	90	H7	GAVTLLVSTVQKREL
43	D6	NQVSYTSWFLDAFNY	91	H6	TLLVSTVQKRELVNP
44	D5	SYTSWFLDAFNYAIL	92	H5	VSTVQKRELVNPASM
45	D4	SWFLDAFNYAILKKI	93	H4	VQKRELVNPASMKQA
46	D3	LDAFNYAILKKIDVL	94	H3	RELVNPASMKQALIA
47	D2	FNAILKKIDVLNLS	95	H2	VNPASMKQALIASAR
48	D1	AILKKIDVLNLSIGG	96	H1	ASMKQALIASARRLP

FIG. 8A

13/13

97	I12	IKDFHVYFRESRDAG
98	I11	DAELHIFRVFTNNQV
99	I10	KQALIASARRLPGVN
100	I9	LIASARRLPGVNMFE
101	I8	SARRLPGVNMFEQGH
102	I7	RLPGVNMFEQGHGKL
103	I6	GVNMFEQGHGKLDLL
104	I5	MFEQGHGKLDLLRAY
105	I4	OGHGKLDLLRAYOIL
106	I3	GKLDLLRAYQILNSY
107	I2	DLLRAYQILNSYKPQ
108	I1	RAYQILNSYKPQASL
109	J12	QILNSYKPQASLSPS
110	J11	NSYKPQASLSPSYID
111	J10	KPQASLSPSYIDLTE
112	J9	ASLSPSYIDLTECPY
113	J8	SPSYIDLTECPYMWP
114	J7	YIDLTECPYMWPYCS
115	J6	LTECPYMWPYCSQPI
116	J5	CPYMWPYCSQPIYYG

FIG. 8B